# How I Learned to Stop Worrying and Love QFT
LMU, Summer 2011

Robert C. Helling
(helling@atdotde.de)
Notes by Constantin Sluka and Mario Flory

August 4, 2011

**Abstract**

Lecture notes of a block course explaining why quantum field theory might be in a better mathematical state than one gets the impression from the typical introduction to the topic. It is explained how to make sense of a perturbative expansion that fails to converge and how to express Feynman loop integrals and their renormalization using the language of distribtions rather than divergent, ill-defined integrals.

## 1 Introduction

Physicists are often lax when it comes to mathematical rigor and use objects that do not exist according to strict mathematical standards or happily exchange limits without justification. This different culture of "everything is allowed as long as it is not proven to be wrong and even then it sometimes ok because we do not actually mean what we are writing" is preferred by many as it allows to "focus on the content rather than the formal aspects" and to progress at a much faster pace.

This attitude can be seen when physicists talk about quantum mechanics and treat operators as if they were matrices and plane waves as if they are elements of the relevant Hilbert space. This is generally accepted since one has the feeling that these arguments can easily be repaired at the expense of clarity by talking about wave packets instead of plane waves and (like it is discussed at length in our "Mathematical Quantum Mechanics" course) by talking about quadratic forms instead of the operators directly.

The situation appears to be very different in the case of quantum field theory: There, most of the time, one deals with perturbative series expansions in the coupling constant without thinking about convergence (or if one spends some thought on this one easily sees that the radius of convergence has to be zero) and the individual terms in the series turn out to be divergent and one obtains reasonable, finite expressions after some very doubtful formal manipulations (often presented as subtracting infinity from infinity in the "right way"). The typical QFT course, unlike quantum mechanics above, does not indicate any way to "repair" these mathematical shortcomings. Often, one is left with the impression that there is some blind faith required on the side of the physicists or at least that some black magic is helping

to obtain numerical values that fit so impressively what is measured in experiments from very doubtful expressions.

In these notes we will indicate some ways in which these treatments can be made more exact mathematically thus providing some cure to the mathematical uneasiness related to quantum field theory. In particular, we will argue that QFT is not “obviously wrong” as claimed by some mistakenly confusing mathematical rigor with correctness.

Concretely, we want to explain how two (mostly independent) crucial steps in QFT can be understood more mathematically:

In a simplified example, we will explore what conclusions can be drawn from the perturbative expansion even though the series does not converge for any finite value of the coupling constant. In particular we will discuss the role of non-perturbative contributions like instantons in the full interacting theory. We will find that up to a certain level of accuracy (depending on the strength of coupling), the first terms of the perturbative expansion do represent the full answer even though summing up all terms leads to infinite, meaningless expressions. Furthermore, at least in principle, using the technique of “Borel resummation” one can express the true expression for all values of the coupling constant in terms of just the perturbative expansion.

As a second step, at each order in perturbation theory, we will see how by correctly using the language of distributions one can set up the calculation of Feynman diagrams without diverging momentum integrals. We will find that these divergences can be understood to arise from trying to multiply distributions. We will set this up as the problem to extend distributions from a subset of all test functions at the expense of a finite number of undetermined quantities that we will identify as the “renormalized coupling constants”. Finally we will understand how these vary when we change regulating functions that were introduced in the procedure which leads to an understanding of the renormalization group in this formalism of “causal perturbation theory”.

The aim is to argue how the techniques of physicists could be embedded in a more mathematical language without actually doing this. At many places we just claim results without proof or argue by analogy (for example we will discuss a one dimensional integral instead of an infinite dimensional path-integral). To really discuss the topic at a mathematical level of rigor requires a lot more work and to large extend still needs to be done for theories of relevance to particle physics.

All this material is not new but well known to experts in the field. Still, we hope that these notes will be a useful complement to standard introductions to quantum field theory for (beginning) practitioners.

## 2 Perturbative expansion — making sense of divergent series

Before we take a look at divergent series, we will first give a brief review of how perturbative expansion is used in quantum field theory.

### 2.1 Brief overview on path integrals

A quantum field theory in Minkowski spacetime is described by a Lagrangian density $\mathcal{L}(\phi, \partial\phi)$ and a generating functional of correlation functions[1]

$$\mathcal{Z}[J] = \int \mathcal{D}\phi e^{i \int \mathrm{d}^4x(\mathcal{L}+J\phi)}. \tag{1}$$

The correlation functions can be obtained by functional derivatives of (1) with respect to $J$.

$$\langle \phi(x_1)\phi(x_2)\dots\phi(x_n)\rangle = \frac{1}{\mathcal{Z}[0]}\left(-i\frac{\delta}{\delta J(x_1)}\right)\left(-i\frac{\delta}{\delta J(x_2)}\right)\dots\left(-i\frac{\delta}{\delta J(x_n)}\right)\mathcal{Z}[J]\Bigg|_{J=0} \tag{2}$$

In this lecture we will use Euclidean signature for the metric instead of Minkowski. The change between the metrics can be performed as rotation of the time axis in the complex plane $t \to -i\tau$ if all expressions are analytic. In Euclidean metric, the exponent in the generating functional is real and falls of at large field values. This gives the path integral a chance to have a mathematical definition in terms of Wiener measures but that will not concern us in these notes.

$$\mathcal{Z}[J] = \int \mathcal{D}\phi e^{\int \mathrm{d}^4x(\mathcal{L}+J\phi)} \tag{3}$$

In general, the integral (3) cannot be computed exactly. For a scalar quantum field theory in Euclidean space the Lagrangian has the form

$$\mathcal{L} = \frac{1}{2}\phi(\Box - m^2)\phi - V(\phi) \tag{4}$$

with $\Box \equiv (\partial_\tau)^2 + (\nabla)^2$.

If the potential $V(\phi)$ vanishes, equation (3) can be formally computed as it becomes an integral of Gaussian type. One therefore arbitrarily splits the Lagrangian into its "kinetic part" $\frac{1}{2}\phi(\Box - m^2)\phi$ and its "interaction part" $-V(\phi)$.

$$\begin{aligned}\mathcal{Z}[J] &= \int \mathcal{D}\phi e^{\frac{1}{2}\int \mathrm{d}^4x\phi(\Box-m^2)\phi} e^{-\int \mathrm{d}^4xV(\phi)} e^{-\int \mathrm{d}^4xJ\phi} \\ &= e^{-\int \mathrm{d}^4xV(\frac{\delta}{\delta J})}\int \mathcal{D}\phi e^{\int \mathrm{d}^4x\frac{1}{2}\phi(\Box-m^2)\phi-J\phi}\end{aligned} \tag{5}$$

To obtain the Gaussian integral one has to complete the square in the exponent. This is achieved by shifting the field $\phi$:

$$\phi' = \phi + (\Box - m^2)^{-1}J \tag{6}$$

[1] This subsection displays some standard expressions to set the context. For many more details see for example [1].

The inverse of $\Box - m^2$, called "Green's function" $G(x-y)$, is a distribution defined by

$$(\Box - m^2)G(x-y) = \delta(x-y). \tag{7}$$

Changing variables in the functional integral (5) leads to

$$\mathcal{Z}[J] = e^{-\int \mathrm{d}^4 x V(\frac{\delta}{\delta J})} \int \mathcal{D}\phi' e^{\int \mathrm{d}^4 x \frac{1}{2}\phi'(\Box - m^2)\phi'} e^{-\int \mathrm{d}^4 x \int \mathrm{d}^4 y \frac{1}{2} J(x) G(x-y) J(y)}. \tag{8}$$

The complicated expression in the middle of equation (8) does not depend on $J$ and will in fact cancel out in equation (2) for the correlation function, so we will just denote it $C$ and forget about it:

$$\mathcal{Z}[J] = Ce^{-\int \mathrm{d}^4 x V(\frac{\delta}{\delta J})} e^{-\int \mathrm{d}^4 x \int \mathrm{d}^4 y \frac{1}{2} J(x) G(x-y) J(y)}$$

Now let us take a look on a specific example for a quantum field theory by choosing a potential for the scalar field. We will consider our favorite $\phi^4$ theory given by the potential

$$V(\phi) = \lambda \phi^4. \tag{9}$$

The next step is to insert this potential in equation (8) and write the exponential as a power series in the coupling strength $\lambda$.

$$\mathcal{Z}[J] = C \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \int \mathrm{d}^4 x_1 \frac{\delta^4}{\delta J(x_1)^4} \dots \int \mathrm{d}^4 x_k \frac{\delta^4}{\delta J(x_k)^4} e^{-\int \mathrm{d}^4 x \int \mathrm{d}^4 y \frac{1}{2} J(x) G(x-y) J(y)} \tag{10}$$

We now found an expression for any general correlation function in terms of an power series expansion in the coupling strength.

$$\langle \phi(y_1) \dots \phi(y_n) \rangle = \frac{\delta}{\delta J(y_1)} \cdots \frac{\delta}{\delta J(y_1)} \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \int \mathrm{d}^4 x_1 \frac{\delta^4}{\delta J(x_1)^4} \cdots \int \mathrm{d}^4 x_k \frac{\delta^4}{\delta J(x_k)^4} \times e^{-\int \mathrm{d}^4 x \int \mathrm{d}^4 y \frac{1}{2} J(x) G(x-y) J(y)} \Bigg|_{J=0} \tag{11}$$

The combinatorics of the occurring expressions in terms of integrals over interaction points $x_i$, Green's functions and external fields can be summarized in terms of Feynman diagrams each standing for a single term in the power series in the coupling constant $\lambda$[2]. In the following, we want to study the convergence behavior of this power series.

## 2.2 Radius of convergence of correlation functions

Let us briefly review the definition of the radius of convergence for a power series from introductory analyis. It is useful to think of a power series to be defined in the complex plane:

$$\sum_{k}^{\infty} \lambda^k (\dots) \quad \lambda \in \mathbb{C} \tag{12}$$

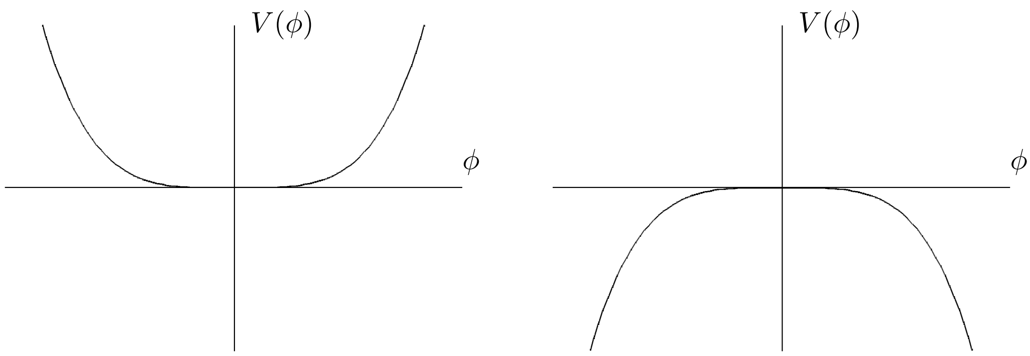
[2]The careful reader wishing to avoid ill-defined expressions using path-integrals, can use this formula as the definition of the terms in the perturbative series.

(If one does not like the idea of a complex coupling strength in a quantum field theory, just restrict to the special $\lambda \in \mathbb{C}$ that happen to be real.). Every power series has a radius of convergence $R \in [0, \infty]$ such that

$$\sum_k^\infty \lambda^k (\dots) \begin{cases} \text{converges} & \forall\, |\lambda| < R \\ \text{diverges} & \forall\, |\lambda| > R. \end{cases} \tag{13}$$

Now we want to find out the radius of convergence for the correlation functions (11) in a quantum field theory. A physicist's argument was given by Freeman Dyson in 1952[2]. Let us take a look on the potential, for example in our $\phi^4$ theory as shown in figure 1. For positive coupling strength $\lambda$ the potential is bounded from below and large values of $\phi$ are strongly disfavored. This behavior, however, gets radically different in case of a negative $\lambda$. The potential becomes unbounded from below and the field $\phi$ will want to run off to $\phi = \pm\infty$. Obviously, such a behavior is highly unphysical, since ever increasing values of $\phi$ would lead to an infinite energy gain. It is thus clear that such a theory cannot lead to healthy correlation functions, in other words for any negative $\lambda$ the power series (12) will diverge[3]. From this we can conclude the radius of convergence being $R = 0$!

$$\sum_k^\infty \lambda^k (\dots) \qquad \text{diverges } \forall \lambda > 0 \tag{14}$$




(a) $\lambda > 0$ (b) $\lambda < 0$

Figure 1: Potentials of $\phi^4$ theory

For readers not satisfied by this argument using physics of unstable potentials for determining the radius of convergence, let us mention an alternative line of argument. Again, consider equation (11), this time, however, we will focus on the Feynman diagrams. At any order $k$ in the perturbation expansion there is a sum of different Feynman diagrams expressing the integrals in (11), where $k$ counts the number of vertices. The combinatorics of all Feynman diagrams shows that the

[3] We expect at least a phase transition when $\lambda$ is changed from positive to negative values.

number of Feynman diagrams grows like $k!$. The power series, therefore, will behave like

$$\sum_{k}^{\infty} \lambda^k k!(\ldots). \tag{15}$$

Assuming that $(\ldots)$ is not surprisingly suppressed for large $k$, the coefficients of $\lambda^k$ grow faster than any power, we again find the radius of convergence $R = 0$.

In the following, we want to give an example, how one can nevertheless make sense of (some) divergent power series.

## 2.3 Non-perturbative corrections

In order to get a feeling for the problem of divergent power series, we will consider a one dimensional toy problem (rather than the infinite dimensional problem of a path integral):

$$\mathcal{Z}(\lambda) = \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2-\lambda x^4} \tag{16}$$

We take $\lambda \geq 0$, so this integral yields some finite, positive number. For $\lambda = 0$ the solution is well known

$$\mathcal{Z}(0) = \sqrt{\pi}. \tag{17}$$

In general, equation (16) can be expressed in terms of special functions, e.g. *Mathematica* gives the solution

$$\mathcal{Z}(\lambda) = \frac{e^{\frac{1}{8\lambda}} K_{1/4}(1/8\lambda)}{2\sqrt{\lambda}} \tag{18}$$

with $K_n(x)$ being the modified Bessel function of the second kind. We call solution (18) the "full, non-perturbative answer". Now we will do the same as in quantum field theory and split the integral into a "kinetic" and an "interaction" part, respectively.

### 2.3.1 Treating the toy model perturbatively

Following the same procedure, we will again expand the "interaction part" $-\lambda x^4$ in a power series:

$$\mathcal{Z}(\lambda) = \int_{-\infty}^{\infty} \mathrm{d}x e^{-x^2-\lambda x^4} = \int_{-\infty}^{\infty} \mathrm{d}x e^{-x^2} \sum_{k=0}^{\infty} \frac{(-\lambda x^4)^k}{k!} \tag{19}$$

Now comes the crucial step and "root of all evil". Following precisely the same steps leading towards equation (11) for correlation functions in quantum field theory, we will change the order of integration and summation, leading to the interpretation of a power series of Feynman diagrams:

$$\mathcal{Z}(\lambda) \text{“} = \text{”} \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \int_{-\infty}^{\infty} \mathrm{d}x\ x^{4k} e^{-x^2} \tag{20}$$

From this step, as we will see later, the problems arise. Although this step is forbidden (as roughly speaking, we are changing the behavior of the integrand at $x =$

$\pm\infty$), we are interested in to what extent a "perturbative solution" obtained from equation (20) will agree with the full, non-perturbative solution (18). Carrying on, we observe that the integral in (20) is now of the type "polynomial times Gaussian" and can be computed with standard methods. We smuggle an addtional factor $a$ into the exponent allowing us to write the integrand as derivatives of $e^{-ax^2}$ with respect to $a$ at the point $a = 1$.

$$\mathcal{Z}(\lambda) = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \int_{-\infty}^{\infty} \mathrm{d}x \frac{\partial^{2k}}{\partial a^{2k}} e^{-ax^2}\Big|_{a=1} = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \frac{\partial^{2k}}{\partial a^{2k}} \sqrt{\frac{\pi}{a}}\Big|_{a=1} \tag{21}$$

Of course we can easily evaluate the derivatives:

$$\frac{\partial^{2k}}{\partial a^{2k}} a^{-\frac{1}{2}}\Big|_{a=1} = \underbrace{\frac{1}{2}\frac{3}{2} \cdot \frac{5}{2}\frac{7}{2} \cdot \frac{9}{2}\frac{11}{2} \cdot \cdots}_{\text{total of } 2k \text{ factors}} \tag{22}$$

In order to find an explicit expression for (22) one can insert factors of 1 between all factors, such that the nominator becomes $(4k)!$:

$$\begin{aligned} \frac{\partial^{2k}}{\partial a^{2k}} a^{-\frac{1}{2}}\Big|_{a=1} &= \underbrace{\frac{1}{2}\frac{2}{2}\frac{3}{2}\frac{4}{4}\frac{5}{2}\frac{6}{6}\frac{7}{2}\frac{8}{8}\frac{9}{2}\frac{10}{10}\frac{11}{2}\frac{12}{12}\cdots}_{\text{total of } 4k \text{ factors}} = \frac{(4k)!}{2^{2k}} \underbrace{\frac{1}{2}\frac{1}{4}\frac{1}{6}\frac{1}{8}\frac{1}{10}\frac{1}{12}\cdots}_{\text{total of } 2k \text{ factors}} \\ &= \frac{(4k)!}{2^{2k}} \frac{1}{2^{2k}(2k)!} = \frac{(4k)!}{2^{4k}(2k)!} \end{aligned} \tag{23}$$

Thus we obtain the "perturbative solution" of problem (16)

$$\mathcal{Z}(\lambda) = \sum_{k=0}^{\infty} \sqrt{\pi} \frac{(-\lambda)^k (4k)!}{2^{4k}(2k)!k!}. \tag{24}$$

Let us take a closer look at this expression. By observing that the denominator of the summand eventually contains smaller factors than the nominator for all $k$ larger than a critical integer, we can realize that the series is divergent. More carefully we can apply Stirling's formula $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ for large values of $k$:

$$\frac{(4k)!}{2^{4k}(2k)!k!} \approx \frac{4^k}{\sqrt{\pi k}} \left(\frac{k}{e}\right)^k \approx \frac{1}{\sqrt{2\pi}} 4^k k! \tag{25}$$

We already know that the sum

$$\sum_{k=0}^{\infty} (-4\lambda)^k k! \tag{26}$$

will diverge. This shows that the power series (24) is divergent and in particular it is not the finite number that we are looking for as an expression for (16).

### 2.3.2 The perturbative and the full solution compared

Even though the perturbative series will diverge, we want to study its numerical usefulness at finite order. After all, one usually computes only a finite number

of Feynman diagrams to obtain only the first few summands of the perturbative expansion. Is there a way to approximate the full, non-perturbative solution (18) from (24)? Let us choose one value for $\lambda$, e.g. $\frac{1}{50}$, and evaluate (18) numerically:

$$\mathcal{Z}\left(\frac{1}{50}\right) = 1.7478812\ldots \tag{27}$$

For the same value of $\lambda$ the evaluation of the first few terms of the infinite sum (24)

$$\mathcal{Z}_N(\lambda) = \sum_{k=0}^{N} \sqrt{\pi}\frac{(-\lambda)^k(4k)!}{2^{4k}(2k)!k!}. \tag{28}$$

gives

$$\mathcal{Z}_5 = 1.7478728\ldots \tag{29a}$$
$$\mathcal{Z}_{10} = 1.7478818\ldots \tag{29b}$$

The first terms of the perturbative solution agree up to six digits! We can use conveniently a figure for plotting higher orders of the perturbative series. Figure 2 shows that the perturbative solution gets in a certain regime very close to the result of the full solution, before the series starts to diverge. We can use a figure as well

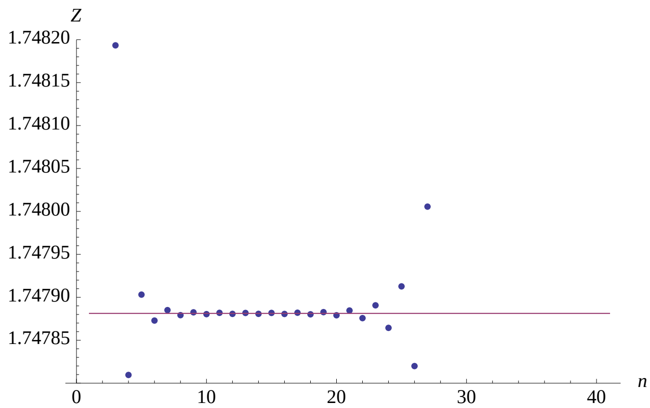



Figure 2: Values of the perturbative series (24) evaluated to order $N$

to compare the solutions for variable $\lambda$. Figure 3 shows nicely the non-perturbative solution and compares it to the perturbative solution for orders of one to twelve. We can see that at some point all approximations given by the perturbative solution will disagree strongly from the full solution!

The question that arises is, how long does the perturbative solution become better before it starts to diverge? Obviously, the fact that it approximates the non-perturbative solution to high precision leads to the great success of quantum field theory, even if for higher orders the series diverges! As we will see now, the perturbative solution (24) is a good approximation as long as we only consider terms up to order $N = O(\frac{1}{\lambda})$. Remembering the dimensionless coupling strength of Quantum Electrodynamics being the Sommerfeld finestructure constant $\alpha \approx \frac{1}{137}$ we can be ensured that perturbation theory will lead to great precision given that the most elaborate QED calculations for $(g-2)$ are to order $N = 7$!

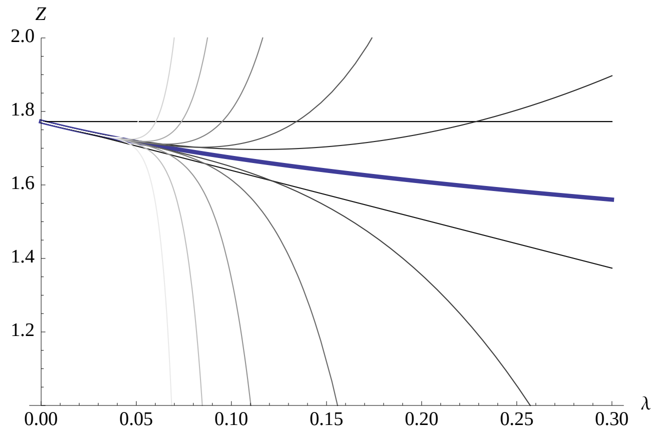



Figure 3: $\mathcal{Z}(\lambda)$ obtained from the full solution (thick) and first approximations from the perturbative series

#### 2.3.3 The method of steepest descend

But what is the origin of the eventual divergence and complete loss of numerical accuracy? It turns out that there are "non-perturbative" terms that do not show up in a Taylor expansion but that become dominant when the perturbative expansion breaks down. To see this, let us substitute $x^2 \equiv \frac{u^2}{\lambda}$ in equation (16):

$$\mathcal{Z}(\lambda) = \frac{1}{\sqrt{\lambda}} \int \mathrm{d}u e^{-\frac{u^2+u^4}{\lambda}} \tag{30}$$

The exponent is strictly negative and its absolute value becomes very large in the limit of small $\lambda$. This allows to perform the method of steepest descent: The main contribution to the integral, as $\lambda \to 0$ comes from the extrema of the integrand

$$u^2 + u^4. \tag{31}$$

In general the method works as follow: For $\Lambda \to \infty$ we want to solve an integral of the general form

$$\int \mathrm{d}x A(x) e^{(i)\phi(x)\Lambda}. \tag{32}$$

One expands now around its extrema[4] $\phi'(x_0) = 0$ and obtains again an integral of "Gaussian times polynomial" type[5]

$$\begin{aligned} &= \sum_{x_0:\phi'(x_0)=0} \int \mathrm{d}x (A(x_0) + (x-x_0)A'(x_0) \dots) e^{\Lambda(\phi(x_0)+(x-x_0)^2\phi''(x_0)+\dots)} \\ &= \sum_{x_0:\phi'(x_0)=0} A(x_0) e^{\Lambda\phi(x_0)} \sqrt{\frac{2\pi}{\phi''(x_0)\Lambda}} \left(1 + O\left(\frac{1}{\Lambda}\right)\right). \end{aligned} \tag{33}$$

[4]Notice that in field theory $\phi'(x) = 0$ is the equation of motion

[5]Corrections from $(x-x_0)A'(x_0)$ can be obtained by doing again the trick of smuggling an $a$ into the exponent and write the term as derivative with respect of $a$ evaluated at $a = 1$.

In our case, the extrema of (31) are $u = 0$ and $u = \pm i/\sqrt{2}$. Expansion around the first yields the perturbative expansion of above. The other two yield contributions like $e^{\frac{1}{4\lambda}}$ that are invisible to a Taylor expansion around $\lambda = 0$, as all derivative vanish here. We have found an example of a "non-perturbative contribution".

The perturbative solution, however, gives meaningful results, as long as its terms are bigger than to the non-perturbative contributions. This allows an estimate, to what order in the perturbative series the expansion around $\lambda = 0$ dominates. This happens also to be the order at which the divergence from the exat solution starts as we are missing the non-perturbative terms:

$$\begin{aligned} e^{-\frac{1}{4\lambda}} &\approx \lambda^k \\ -\frac{1}{4\lambda} &\approx k \cdot \ln(\lambda) \\ k &\approx \frac{1}{\lambda} \end{aligned} \tag{34}$$

We have seen that the perturbative analysis of (30) requires as well expansions around the other extrema, besides $\lambda = 0$! Combining all power series together, the resulting perturbative solution has a chance to converge. Before we continue to the mathematical discussion of the problem how finite results can be obtained from divergent series, we will take a look on examples of non-perturbative contributions in physics.

### 2.3.4 Instantons

"Field configurations" contributing $e^{-\frac{1}{\sim\lambda}}$ are called "instantons". Usually these contributions are hard to calculate, in some situations, however, one can find the result. Consider for example a gauge theory[6]

$$S = \int \mathcal{L} = \int \frac{1}{g^2} \operatorname{tr}(F \wedge *F) \tag{35}$$

The stationary point we use for the expansion is given by the equations of motion

$$\mathrm{d}F = 0 \tag{36a}$$

$$\mathrm{d} * F = 0 \tag{36b}$$

The first equation (36a) is automatically fulfilled once we express the field-strength in terms of a vector potential $F = \mathrm{d}A$. The second (36b) is automatically solved if it happens that

$$F = *F. \tag{37}$$

One calls solutions to (37) instantons. In terms of the vector potential $A$ (37) is a first order partial differential equation as compared to (36b) which is second order. One can easily see that there exist no solution in Lorentzian metric as the Hodge star squares to $-1$ on 2-forms.

$$F = *F = * * F = -F. \tag{38}$$

In Euclidean metric, however, such solutions exist because of $* * F = F$. As it turns out (as one can for example argue using the Atiya-Singer index theorem), for

[6] Hodge $*$ operator: $*F^{\mu\nu} \equiv \epsilon^{\mu\nu\sigma\tau} F_{\sigma\tau}$. More details can be found in chapters 1.10 and 10.5 of [3].

a compact manifold $M$, the action in the instanton case yields an integer (up to a pre-factor):

$$\int_M \mathrm{tr}(F \wedge F) \in 8\pi^2\mathbb{Z} \tag{39}$$

This leads to

$$e^{\frac{1}{g^2}\int \mathrm{tr}(F\wedge *F)} = e^{\frac{1}{g^2}\int \mathrm{tr}(F\wedge F)} = e^{-\frac{1}{g^2}8\pi^2 N} \tag{40}$$

### 2.3.5 Dual theories

Sometimes a quantum field theory with coupling constant $\lambda$ can be rewritten in terms of another (or possibly the same) quantum field theory with coupling $\tilde{\lambda} = 1/\lambda$. One calls such a relation between two theories a "duality". In many examples, such theories arise from string theory constructions, where the coupling $\lambda$ can be given a geometric meaning. Imagine for example a problem of a quantum field theory on a torus. A torus can be viewed as $\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$ with $\tau \in \mathbb{C}\backslash\mathbb{R}$. The torus has a basis of two non-contractible circles, one that goes along the real axis from 0 to 1 and one that goes from 0 to $\tau$. This choice of basis, however, is not unique: For example, swapping these cycles corresponds to a substitution $\tau \to -1/\tau$. If the torus parameter $\tau$ is identified with the coupling strength a duality has been found since both $\tau$ and $-1/\tau$ describe geometricaly the same torus! To make contact with our discussion above, we should identify $\lambda$ with the imaginary part of $\tau$. The duality allows for a Taylor expansion of the non-perturbative contributions via

$$e^{-\frac{1}{4\lambda}} = e^{-\frac{\tilde{\lambda}}{4}} = \sum_k^\infty \frac{\left(-\tilde{\lambda}\right)^k}{k!4^k} \tag{41}$$

Troublesome terms in one theory are therefore perfectly defined in the dual theory. The caveat however is the difficulty of actually proving that $\lambda \to 1/\lambda$ is a symmetry of the quantum field theory at hand.

## 2.4 Asymptotic series and Borel summation

In the following, we take a look on the mathematical situation of asymptotic series. This discussion is based on chapter $XII$ of [4].

**Definition 1** *Let* $f\colon \mathbb{R}_{\geq 0} \to \mathbb{C}$*. The series* $\sum_n^\infty a_n z^n$ *is called* asymptotic *to* $f$ *as* $z \searrow 0$ *iff*

$$\forall N \in \mathbb{N} : \lim_{z\searrow 0} \frac{f(z) - \sum_n^N a_n z^n}{z^n} = 0 \tag{42}$$

*For* $z \in \mathbb{C}$ *a analog definition is possible.*

Obviously, every function can have at most one asymptotic expansion. This can be seen by assuming two asymptotic expansions $a_n$ and $\tilde{a}_n$. (42) requires that $a_n = \tilde{a}_n$. Otherwise, let $n$ be the smallest index for which $a_n \neq \tilde{a}_n$ and

$$\lim_{z\searrow 0} \frac{\sum_k (a_k - \tilde{a}_k) z^k}{z^n} = a_n - \tilde{a}_n \stackrel{!}{=} 0. \tag{43}$$

The other way around is not true, as can be seen by $f(z) = e^{-\frac{1}{z}}$ and $\tilde{f}(z) = 0$ having both the asymptotic series $\sum_k^\infty 0 \cdot z^k$. This means that knowing the asymptotic series

of a function tells us nothing about $f(z)$ for a non vanishing $z$, we only know how $f(z)$ approaches $f(0)$ as $z \searrow 0$.

We try to find a stronger definition of an asymptotic series, allowing us to uniquely recover one function. The following theorem helps us to find the necessary conditon:

**Theorem 2** ***(Carleman's theorem)*** *Let $g$ be an analytic function in the interior of $S = \{z \in \mathbb{C} | \, |z| \leq B, |\arg\, z| \leq \frac{\pi}{2}\}$ and continous on $S$. If for all $n \in \mathbb{N}$ and $z \in S$ we have $|g(z)| \leq b_n |z|^n$ and $\sum_n^\infty b_n^{-\frac{1}{n}} = \infty$, then $g$ is identically zero.*

*A simpler special case of the theorem is found by considering $g$ an analytic function in the interior of $S_\epsilon = \{z \in \mathbb{C} | \, |z| \leq R, |\arg\, z| \leq \frac{\pi}{2} + \epsilon\}$ for some $\epsilon > 0$ and continous on $S_\epsilon$. If there exist $C$ and $B$ so that $|g(z)| < CB^n n! \, |z|^n \;\; \forall z \in S$ and $\forall n$, then $g$ is identically zero.*

In order to find a unique function for an asymptotic series, we use Carleman's theorem to define "strong asymptotic series".

**Definition 3** *Let $f$ be an analytic function on the interior of $S_\epsilon = \{z \in \mathbb{C} | \, |z| \leq R, |\arg\, z| \leq \frac{\pi}{2} + \epsilon\} \to \mathbb{R}$. The series $\sum_n^\infty a_n z^n$ is a* strong asymptotic series *if there exist $C, \sigma$ so that $\forall N \in \mathbb{N}, z \in S_\epsilon$ the* strong asymptotic condition

$$\left| f(z) - \sum_n^N a_n z^n \right| \leq C\sigma^{N+1}(N+1)! \, |z|^{N+1} \tag{44}$$

*is fulfilled.*

This means, if we are given a strong asymptotic series, we can recover by theorem 2 the function! Assume for example $\sum_n^\infty a_n z^n$ is a strong asymptotic series for two functions $f$ and $g$, respectively. Then

$$|f(z) - g(z)| \leq 2C\sigma^{N+1}(N+1)! \, |z|^{N+1} \Rightarrow f = g \tag{45}$$

The strong asymptotic condition (44) implies $|a_n| \leq C\sigma^n n!$. This is precisely the growth behavior of (24) we found in our toy example, where $C = \frac{1}{\sqrt{2\pi}}$ and $\sigma = 4$. The necessary conditions, therefore, are fulfilled in our toy model (assuming analyticity away from 0 of course).

By now, we learned that a strong asymptotic series (in particular the type we obtain in quantum field theory) although not converging has the chance to be a unique approximation to one function. The final question is, how one can obtain this function $f$ from its strong asymptotic series. In the last theorem we introduce the method of "Borel summation" to obtain a final result. We can define a convergent series by taking out a factor of $n!$ from the coefficients:

**Theorem 4** ***(Watson's theorem)*** *If $f : S_\epsilon \to \mathbb{R}$ has a strong asymptotic series $\sum_n^\infty a_n z^n$, we define the* Borel transform

$$g(z) = \sum_n^\infty \frac{a_n}{n!} z^n. \tag{46}$$

*The Borel transform converges for $|z| < \frac{1}{|\sigma|}$. We obtained a convergent power series with finite radius of convergence, which, as it turns out, can be analytically continued*

*to all complex $z \in \mathbb{C}$ with $|\arg\, z| < \epsilon$. Then the function $f$ is given by the Laplace transform*

$$f(z) = \int_0^\infty \mathrm{d}b\ g(bz)e^{-b}. \tag{47}$$

This Laplace transform is called “inverse Borel transform” and the method outlined here is known as “Borel summability method”. It describes how to obtain a finite answer from divergent series, that is formally a sum for the series.

Let us make a sanity check. Using $\int_0^\infty \mathrm{d}x\ x^k e^{-x} = k!$ we can plug the definition of the Borel transform into (47), formally interchange the sum and the integration and obtain

$$f(z) = \int_0^\infty \mathrm{d}b\ g(bz)e^{-b} = \int_0^\infty \mathrm{d}b \sum_n \frac{a_n}{n!} b^n z^n e^{-b} \text{ “=” } \sum_n a_n z^n. \tag{48}$$

So at least for analytic functions we do recover the original function.

## 2.5 Summary

We have learned why for $N < \mathcal{O}(\frac{1}{\lambda})$ the sum of the first $N$ terms of the perturbation expansion is numerically good, even when the original series $\sum_n a_n z^n = \infty$ diverges. This way we approximate the true function up to instantonic terms of the order of $e^{1/\lambda}$ which a Taylor expansion cannot resolve. Given that the coefficients $a_n$ obey the strong asymptotic condition $|a_n| \leq C\sigma^n n!$, which is usually the case when using Feynman diagrams, the Borel transform exists and one can compute the Borel summation. Unfortunately this is more a theoretical assurance that perturbation theory can be given a mathematical meaning even though it does not converge, since in order to really compute the integral (47) one has to know the analytic continuation of $g$ which requires knowledge of *all* coefficients $a_n$ and not just the first $N$.

# 3 Regularization and renormalization as extensions of distributions

In the previous section, we learned how to make sense of (some) divergent series of the form $\sum_{k=0}^\infty a_k = \infty$, but in QFT the factors $a_k$ are typically complicated mathematical expressions described by Feynman diagramms, and generically, these expressions diverge themselves, creating a need for renormalization techniques.
A typical example of a divergent diagramm (in 4 dimensions) is shown in Figure 4.

The term described by this diagramm reads (without unimportant factors):

$$\int d^4k \frac{1}{k^2+m^2} \frac{1}{(p_1+p_2+k)^2+m^2} \xrightarrow{k \gg p,m} \int \frac{k^3 dk}{k^4} = \infty$$

where $m$ denotes the mass of the scalar particles we are scattering. This integral obviously diverges logarithmically for $k \to \infty$ as shown above. The most straightforward approach to this problem is to introduce a cut-off energy-scale $\Lambda$, such that the divergence at the upper boundary becomes

$$\int^\Lambda \frac{k^3 dk}{k^4} \sim \log(\Lambda),$$

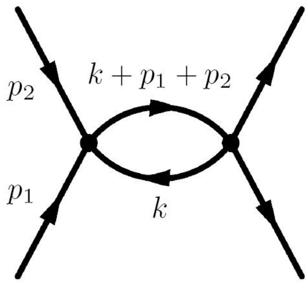



Figure 4: Divergent 1-loop diagramm

Usually, such blunt cut-off regularization is incompatible with the symmetries of the theory at hand and is thus only useful to estimate "how divergent" a diagram is (a notion we will below formalize as the "singular degree") and has to be replaced by more sophisticated methods like dimensional or Pauli-Villars regularization in more practical applications.

In these notes, instead of momentum representation, we will work in position space where instead of loop momenta one integrates over the position of the interaction vertices.

What was $1/(k^2 + m^2)$, is now the propagator $G$ defined by the equation

$$(\Box + m^2)G(x) = \delta(x), \tag{49}$$

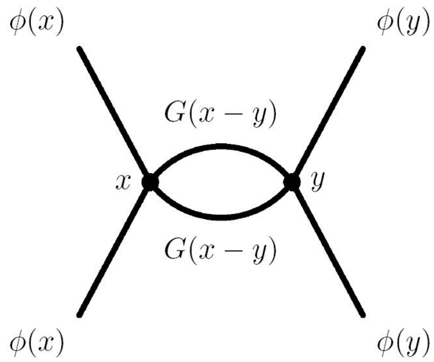



Figure 5: Divergent 1-loop diagramm in position-space language

we can compute the same diagramm in position space language, which then reads (see Figure 5):

$$\int d^4x \int d^4y \; \phi_0^2(x) G^2(x-y) \phi_0^2(y) = \int d^4x \int d^4u \; \phi_0^2(x) G^2(u) \phi_0^2(x-u) \tag{50}$$

The approach of "causal perturbation theory" or "Epstein-Glaser regularization" is to take seriously the fact that the propagator is really a distribution and in the above expression, we are trying to multiply distributions which in general is

undefined. This approach is advocated in the book by Scharf [5]. Here, we will follow (a simplified, flat space version) of [6] and in particular [7].

Specifically, in the defining equation (49) $\delta$ is not a *function* but a *distribution* (physicists writing $\delta(x)$ are trying to imply that this is the kernel of the distribution $\delta$, i.e. that $\delta$ arises by muliplying the testfunction by a function $\delta(x)$ and then integrating over $x$, which of course does not exist). Thus, we should interpret $G(x)$ as a distribution as well (a priori it is only a weak solution of the differential equation (49)). But as it is in general not possible to multiply distributions, as we will see later, we do not have a naive way to obtain "$G^2$" as a distribution. In this chapter, our goal will be to understand renormalization techniques in terms of distributions. Our route will be led by the question how to define the product of two distributions that are almost everywhere functions (which can be multiplied). We will first, therefore, recapitulate what distributions actually are. Then, we will see in which cases it is possible to multiply distributions and in which it is not. This will lead us to the renormalization techniques we are searching for.

## 3.1 Recapitulation of distributions

Distributions are generalized functions. Like in many other cases of generalizations this is done via dualization: Starting from an ordinatry function $f$ (in our case locally integrable, that is $f\colon \mathbb{R}^n \to \mathbb{C}$ with $\int_K |f| < \infty$ for each compact $K \subset \mathbb{R}^n$, so that "divergence of the integral $\int |f|$ at infinity" is tolerated) one can view it as a linear functional $T_f$ (called a "regular distribution") on the functions of compact support via

$$T_f\colon \phi \mapsto \int f\phi. \tag{51}$$

As the map $f \mapsto T_f$ is injective we can use the $T_f$'s to distinguish the different $f$'s and view $T_f$ in place of $f$. This suggests to generalize the construction to all linear functionals $T : \phi \mapsto T(\phi)$ called distributions of which the regular ones arising from functions $f$ as above are a subset.

Specifically, distributions are defined to be linear and continuous functionals on the space of *test functions* $D(\mathbb{R}^n) = \mathcal{C}_0^\infty(\mathbb{R}^n)$ (the subscript 0 meaning compact support) equipped with an appropriate topology that will not concern us here. So formally, we can denote the distributions to be elements of a space:

$$D'(\mathbb{R}^n) = \{T : D(\mathbb{R}^n) \to \mathbb{C} \mid T \text{ is linear and continuous}\}$$

Besides the regular distibutions $T_f$ encountered above (of which the function $f$ is called the *kernel*) the typical example of a *singular distribution* is the $\delta$-distribution: if we take a given test function $\phi(x) \in D$, then the $\delta$-distribution is defined to be the functional $\delta[\phi] = \phi(0)$. This distribution is not regular even though physicists pretend it to be with a kernel $\delta(x)$ that is so singular at $x = 0$ that $\int \delta(x) = 1$ even though it vanishes for all $x \neq 0$.

Later, we will make use of the fact that distributions can be differentiated. Using integration by parts in the integral representation of a regular distribution, we easily obtain $T_{f'}[\phi] = -T_f[\phi']$ which enables us to define the derivative of a distribution to be $T'[\phi] \equiv -T[\phi']$. Thus we can take the derivative of a regular distribution $T_f$ even if the kernel $f$ is not differentiable.

The only operation defined on functions that does not directly carry over to distributions is (pointwise) multiplication $(f \cdot g)(x) = f(x)g(x)$. Already $L^1_{loc}$ is not

closed under multiplication (recall that in order for a function to be in $L^1_{loc}$ it must not have singularities that go like $1/x^\alpha$ with $\alpha \geq 1$, a property not stable under multiplication) and in general the product of distributions is not defined. Of course, as long as with $f$ and $g$ also $f \cdot g \in L^1_{loc}$ we still have the regular distribution $T_{f \cdot g}$ and, from a technical perspective, in this sections, we will deal with the problem to extend a distribution that can be written as $T_{f \cdot g}$ for a subset of test functions (those that vanish where $f \cdot g$ is too singular to be in $L^1_{loc}$) to all test functions.

To this end, for any distribution $T \in D'$ we define the *singular support* of T (singsupp($T$)) as the smallest closed set in $\mathbb{R}^n$ such that there exists a function $f \in L^1_{\text{loc}}$ with $T[\phi] = T_f[\phi]$ for all $\phi \in D$ with $\text{supp}\phi \cap \text{singsupp}(T) = \emptyset$. For example $\text{singsupp}(\delta) = \{0\}$, and the corresponding function $f \in L^1_{\text{loc}}$ is simply $f(x) = 0$. So the idea behind this definition is that every distribution can be written as regular distributions as long as it is only applied to test functions which vanish in a neighbourhood of the distribution's singular support, which enables us to multiply distributions if we manage to take care of the singular support.

## 3.2 Definition of $G^2$ in $D'(\mathbb{R}^4\backslash\{0\})$

Coming back to our concrete field-theoretic problem for a moment, we now have to answer one important question: what is the singular support of $G$? Here, we can utilize the fact that $\Box + m^2$ is an *elliptic* operator[7] and it can be shown that if two distributions $T$ and $S$ are related by $\sigma T = S$ with an elliptic operator $\sigma$, then $\text{singsupp}(T) \subset \text{singsupp}(S)$, so we immediately see $\text{singsupp}(G) \subset \{0\}$.

It is clear that the singularity of $G(x)$ at $x = 0$ corresponds to the divergence of the momentum-integral at high energies $\Lambda \to \infty$, because in order to probe small distances, short wavelengths which correspond to high momenta are needed, therefore we speak of UV-divergencies. In this regime, we can set $m^2 \approx 0$, and so (49) simplifies to $\Box G(x) = \delta(x) \Rightarrow G(x) \sim \frac{1}{x^2}$ for small $x$.[8]

Using this, we see that the position-space integral $\int_{|x|>\frac{1}{\Lambda}} d^4x\, G^2(x)$ again diverges as $\log(\Lambda)$. Because of this divergence $G^2(x) \notin L^1_{\text{loc}}(\mathbb{R}^4)$, so $G^2$ is still not defined as distribution in $D'(\mathbf{R}^4)$. Nevertheless, we can use $G^2(x)$ as kernel of a distribution in $D'(\mathbf{R}^4\backslash\{0\}) = \big\{T : D(\mathbb{R}^4\backslash\{0\}) \to \mathbb{C} \mid \text{linear and continuous}\big\}$ where $D(\mathbb{R}^4\backslash\{0\})$ is the set of test functions with $\{0\} \notin \text{supp}\phi$.

We now managed to define a distribution $G^2$, but we still have to extend it from $D(\mathbb{R}^4\backslash\{0\})$ to $D(\mathbb{R}^4)$. Formally, as a linear map, we have to say what values the extension takes on $D(\mathbb{R}^4)/D(\mathbb{R}^4 \setminus \{0\})$ which is still an infinite dimensional vector space. To control this infinity, we will use the *scaling degree.*

## 3.3 The scaling degree and extensions of distributions

Consider a scaling-map $\Lambda$ acting on test functions:

$$\begin{aligned} \mathbb{R}_{>0} \times D(\mathbb{R}^n) &\to D(\mathbb{R}^n) \\ (\lambda,\phi) &\mapsto \phi_\lambda(x) \equiv \lambda^{-n}\phi(\lambda^{-1}x) \end{aligned}$$

[7]Explaining it without going into details, a differential operator which is defined as polynomial of $\vec{\partial}$ (with possible coordinate dependent coefficients) is elliptic if it is non-zero if we replace $\vec{\partial}$ with any non-zero vector $\vec{y}$. In our euclidian examples, $\Box = \sum_{i=1}^4 \partial_i^2 \to |\vec{y}|^2 > 0$ foy any non-zero $\vec{y}$

[8]The relation $\delta(\vec{x}) = -\frac{1}{4\pi}\Box\frac{1}{|\vec{x}|}$ is well known to hold in 3 dimensions. In general, $\Box|x|^{2-n} \propto \delta$ in $n$ dimensions

The pullback of this map to the space distributions reads

$$(\Lambda^* T)[\phi] = T[\phi_\lambda] \equiv T_\lambda[\phi],$$

which for regular distributions gives

$$T_{f,\lambda}[\phi] = \int \frac{d^n x}{\lambda^n} f(x)\phi\left(\frac{x}{\lambda}\right) = \int d^n x f(\lambda x)\phi(x),$$

so power of $\lambda$ in the scaling map acting on test functions is chosen such that the kernel $f$ transforms in a simple manner without prefactor. We now define the *scaling degree* ($sd$) of $T \in D'(M \subset \mathbb{R}^n)$:

$$sd(T) = \inf\left\{\omega \in \mathbb{R} \middle| \lim_{\lambda \searrow 0} \lambda^\omega T_\lambda = 0\right\}$$

To understand this definition, we have to note several properties:

- $sd(T) \in [-\infty, \infty[$
- For regular distributions $sd(T_f) \leq 0$
- $sd(\delta) = n$
- $sd(\partial^\alpha T) \leq sd(T) + |\alpha|$ with some multi-index $\alpha$
- $sd(x^\alpha T) \leq sd(T) - |\alpha|$ with some multi-index $\alpha$[9]
- $sd(T_1 + T_2) = \max\{sd(T_1), sd(T_2)\}$

This leads us to the following important theorem:

**Theorem 5** *If $T_0 \in D'(\mathbb{R}^n \backslash \{0\})$ is a distribution with $sd(T_0) < n$, then there is a unique distribution $T \in D'(\mathbb{R}^n)$ with $sd(T) = sd(T_0)$ extending $T_0$.*

The proof of uniqueness is quite easy: We do it by assuming the existence of two solutions $T$ and $\tilde{T}$ extending $T_0$, and showing a contradiction. Obviously supp$(T - \tilde{T}) = \{0\}$ and from this it follows that $T - \tilde{T} = P(\partial)\delta$ with some polynomial $P$. As can be seen from the above notes, $sd(P(\partial)\delta) \geq n$ and this would be a contradiction to $sd(T) = sd(\tilde{T}) = sd(T_0) < n$. Existence is shown constructively using a smooth cut-off function $c_\epsilon(x)$ that is 1 outside a ball of radius $2\epsilon$ and and vanishes in a ball of radius $\epsilon$. Then we can define

$$T[\phi] = \lim_{\epsilon \searrow 0} T_0[c_\epsilon \phi], \tag{52}$$

where one still has to show that the above limit exists in the sense of distributions.

The theorem above now enables us to uniquely extend distributions of low scaling degree to the full space $D'(\mathbb{R}^n)$, but what about distributions with scaling degree $\geq n$? We will solve this problem in the next section, and afterwards we will be able to return to our field-theoretic problem of understanding the nature of $G^2$.

But first we have to determine what the scaling degree of the massive propagator $G$, defined by $\delta = (\Box + m^2)G$. We know that $sd(\delta) = n$, and therefore $sd((\Box + m^2)G) = n$ too. If we denote $sd(G)$ by $w$, from the above items it follows that $sd(\Box G) = w + 2$, $sd(m^2 G) = w$ and therefore $sd((\Box + m^2)G) = w + 2$. From this it follows that $w = n - 2$ even for the massive propagator.

[9]Remember that distributions do not depend on coordinates, only their kernels. Here we used the definition $(x^\alpha T)[\phi] \equiv T[x^\alpha \phi]$

### 3.4 Case of distributions with high scaling degree

Considering now a distribution $T_0 \in D'(\mathbb{R}^n\backslash\{0\})$ with $sd(T_0) \geq n$, uniqueness as in the above theorem does not hold anymore. But if we take a test function $\phi \in D(\mathbb{R})$ which vanishes of order $\omega \equiv sd(T_0) - n$ ("*singular order*") at $x = 0$, i.e. which can be written as $\phi(x) = \sum_{|\alpha|=\lfloor\omega\rfloor+1} x^\alpha \phi_\alpha(x)$ where $\phi_\alpha(0)$ is finite and $\lfloor\omega\rfloor$ denotes the largest integer not bigger than $\omega$, we can define $T[\phi] \equiv \sum_{|\alpha|=\lfloor\omega\rfloor+1}(x^\alpha T_0)[\phi_\alpha]$. Then the distribution $x^\alpha T_0$ has scaling degree less than $n$ and can thus be uniquely extended.

A general test function can of course be written as a sum of a function vanishing of order $\omega$ and a polynomial of degree at most $\omega$ by subtracting and adding the order $\omega$ Taylor polynomial at $x = 0$:

$$\phi_s(x) \equiv \phi(x) - \sum_{|\alpha|\leq\omega} \frac{x^\alpha}{|\alpha|!}\partial^\alpha\phi(0) \tag{53}$$

This procedure of subtracting the terms leading to divergencies is the *regularization* in this framework. Since the extended distribution $T$ beeing applied to $\phi_s$is unique, by linearity, we still have to define $T$ only on the monomials in $x$ of maximal degree $\omega$. There is no further restriction on doing this and this ambiguity in the extension $T$ is what one would have expected: Changing the value of $T$ on a monomial $x^\alpha$ correspond to adding a multiple of $\partial^\alpha\delta$ to $T$.

Note well that the arbitratry values of $T[x^\alpha]$ are exactly those where $T_0[x^\alpha]$ was undefined (divergent in physicists' parlance) and selecting a certain value corresponds to picking a *counter term.* procedure known as *renormalization*, as formally infinite values are replaced by finte ones (that have to be fixed by further physical input like the measurement of the "physical mass" or the physical "coupling constant"). In the following small sections, we will try out this method in a few easy, concrete examples.

#### 3.4.1 Example in $n = 1$

In order to let our steps so far become clearer, we are going to apply them to a simple example in $n = 1$. In fact, this example shows already the full regularization and renormalization procedure.

As can be easily checked, the function $f(x) = \frac{1}{|x|}$ is not an element of $L^1_{loc}(\mathbb{R})$ because of its pole at $x = 0$ is not integrable (it is of course log-divergent), so we can not a priori use it as kernel of a distribution $T_f \in D'(\mathbb{R})$ as we have seen in section 3.1. But $f(x) \in L^1_{loc}(\mathbb{R}\backslash\{0\})$ and $sd(T_f) = 1 = n$, therefore $\omega = 0$. This means that for a test function $\phi(x)$ with $\phi(0) = 0$ we can define $T_f[\phi] \equiv \int dx \frac{\phi(x)}{|x|}$ which gives a finite result: Using l'Hôpital's rule, we see $\lim_{x\to\pm 0} \frac{\phi(x)}{|x|} = \lim_{x\to\pm 0} \frac{\phi'(x)}{\mathrm{sign}(x)} =$ finite and thus the integrand is finite everywhere. This is similar to what we have done in sections 3.2 and 3.3. For other test functions, we can again (as in this section above) define $\phi_s(x) \equiv \phi(x) - \phi(0)$. Afterwards, we write the general extension for a distribution acting on $\phi$ as $T_f[\phi] \equiv T_f[\phi_s] + c\phi(0)$ with one arbitrary constant $c$ of our choice.

The careful reader will have realised that there is still a problem as $\phi_s$ fails to have compact support when $\phi(0) \neq 0$ and thus the integration now diverges at the boundary $x \to \pm\infty$. We will deal with this problem below but the important observation is that the divergence in the ultraviolet, that is at small $x$ is cured.

### 3.4.2 Example in $n = 4$

In our field theoretic problem (50) from above, we have $G^2 \sim \frac{1}{x^4}$ in $\mathbb{R}^4$ which is quite similar to the previous example, as it is the kernel of a distribution $T_{G^2} = T_{\frac{1}{x^4}} \in D'(\mathbb{R}^4 \setminus \{0\})$. Again, we are looking for an extension. Once more, we have $sd(G^2) = 4 = n$. Regularization and renormalization are as in the example above and yield

$$T^r_{\frac{1}{x^4}}[\phi] = \int d^4x \frac{\phi(x) - \phi(0)}{x^4} + c\delta[\phi] \tag{54}$$

with $T^r_{\frac{1}{x^4}} \in D'(\mathbb{R}^4)$ and arbitrary $c$. Again, we successfully got rid of the problems at $x = 0$ (at the cost of introducing one constant $c$).

This concludes our calculation of the fish diagram Fig. 4 that computes a contribution of the form $\phi(x)^2\phi(y)^2$ to the effective action of the theory. Since the ambiguous term we found is $c\delta(x-y)$, the ambiguity in the effective action is indeed $\phi(x)^4\delta(x-y)$. We see, that it corresponds to the counter term Fig. 6 and renormalizes the coupling constant (the coefficient of the $\phi^4$-term in the action).

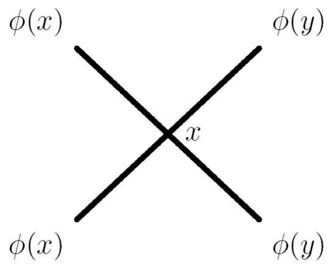



Figure 6: Counter-term diagramm.

### 3.4.3 Example with $sd(T) > n$ and preservation of symmetry

In the two examples above, we had both times distributions $T$ with $sd(T) = n$ which led us to the introduction of one arbitrary constant $c$. This amount of ambiguity increases with $sd(T)$, but not all possible polynomials $P(\partial)\delta$ allowed by the counting can arise physically. In particular, we require that our theory is still Lorentz invariant after renormalization and if it has a local gauge symmetry before that needs to be maintained as well (otherwise one has an *anomaly* that renders the theory ill-defined at the quantum level since the number of degrees of freedom changes upon renormalization).

Let us consider one example where SO(4) invariance (the euclidian version of the Lorentz group SO(3,1)) selects a subset of the possible counter terms.

In a theory with potential $\propto \phi^4$ (*quartic interaction*), there cannot only be diagramms like Figure 4, but also such ones like Figure 7, known as the *setting sun* diagram.

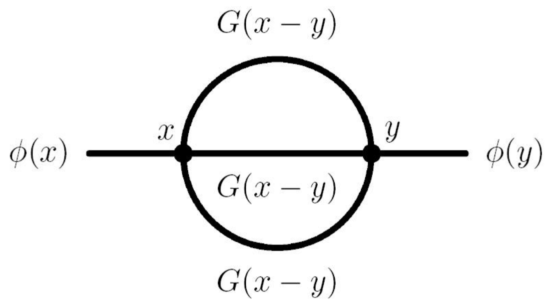



Figure 7: Setting sun diagramm in quartic interaction

The term encoded by this diagramm obviously contains $G^3 \sim \frac{1}{x^6}$, which has $sd(G^3) = 6 > n = 4$. By performing the same steps as above, in this case we a priori get an ambiguity $c_1\delta + c_2^i\partial_i\delta + c_3^{ij}\partial_i\partial_j\delta$ with in total $1 + 4 + \frac{4(4+1)}{2} = 15$ arbitrary constants, but upon imposing SO(4)-invariance this reduces to $c_1\delta + c_3\Delta\delta$ with only 2 arbitrary constants.

In the effective action, as above, they contribute to the quadratic terms (as the diagram Fig. 7 has two external lines) $\phi(x)(c_1\delta(x-y) + c_3\Delta\delta(x-y))\phi(y) = \phi(x)(c_3\Delta + c_1)\phi(x)\delta(x-y)$. We recognize that $c_3$ is a *wave function renormalization* while $c_1$ renormalizes the mass-term $m^2\phi^2$.

The fact that $\phi^4$-theory is renormalizable in $n = 4$ means that these two counter terms and the one in the previous subsection are the only ambiguities that arise when any Feynman diagram of the theory is renormalized, a proof of being well beyond the scope of these notes.

### 3.5 Regaining compact support and RG flow

In the above calculations, we ignored an important problem: $\phi_s(x) = \phi(x) - \phi(0)$ is not necessarily a test function, as it obviously has $\lim_{x\to\infty} = -\phi(0)$, therefore for example the integral $\int dx \frac{\phi(x)-\phi(0)}{|x|}$ that we encountered in section 3.4.1 may diverge at infinity. We can solve this by introducing a function $w(x) \in D(\mathbb{R}^n)$ with (without loss of generality) $w(0) = 1$. We then change the regularized part (i.e. the part without arbitrary constants) of the integral in (54) to

$$T_{\frac{1}{|x|}}[\phi] \equiv \int dx \frac{\phi(x) - w(x)\frac{\phi(0)}{w(0)}}{|x|}. \tag{55}$$

This is a special case of the general formula

$$\phi_s(x) \equiv \phi(x) - \sum_{|\alpha|\leq\omega} \frac{x^\alpha w(x)}{|\alpha|!} \left(\partial^\alpha \frac{\phi(x)}{w(x)}\right)\Bigg|_{x=0},$$

which replaces equation (53). Starting from (55) we can now write

$$T_{\frac{1}{|x|}}[\phi] = \int dx \frac{w(x)(\phi(x) - \phi(0))}{|x|} + \underbrace{\int dx \frac{(1-w(x))\phi(x)}{|x|}}_{=T_{\frac{1-w(x)}{|x|}}[\phi]}.$$

The second term already is a perfectly fine distribution, the first term can be manipulated in the following way:

$$\int dx \frac{w(x)(\phi(x) - \phi(0))}{|x|} = \int dx \frac{w(x)}{|x|} \int_0^x du \phi'(u)$$

Now, in the inner integral we can substitute $u = tx$ and afterwards interchange the integrals:

$$\int dx \frac{w(x)}{|x|} \int_0^1 dt x \phi'(tx) = \int_0^1 dt \int dx \,\mathrm{sign}(x) w(x) \phi'(tx)$$

After this, we can substitute $y = tx$ in the inner integral, giving us

$$\begin{aligned}
&\int_0^1 dt \int \frac{dy}{t} w\left(\frac{y}{t}\right) \mathrm{sign}\left(\frac{y}{t}\right) \phi'(y) \\
&= \int dy \underbrace{\int_0^1 \frac{dt}{t} w\left(\frac{y}{t}\right) \mathrm{sign}(y)}_{\equiv f(y)} \phi'(y) \\
&= T_f[\partial\phi] = -(\partial T_f)[\phi].
\end{aligned}$$

So also the first term is a good distribution. The function $f(y)$ defined above as function of $y$ is well behaved, as $w(x)$ which enters its definition is a test function, and therefore extremely well behaved, in particular vanishes for large arguments, so taking $t \to 0$ does not introduce problems.

As an example, let us now set $w(x) = \theta(1 - M|x|)$ (or actually a smoothed out version of this non-continuous function)

$$\begin{aligned}
f(y) &= \int_0^1 \frac{dt}{t} \theta\left(1 - M\frac{|y|}{t}\right) \mathrm{sign}(y) \\
&= \int_{M|y|}^1 \frac{dt}{t} \theta\left(1 - M|y|\right) \mathrm{sign}(y) \\
&= -\ln(M|y|) \theta\left(1 - M|y|\right) \mathrm{sign}(y)
\end{aligned}$$

Morally, we regularized our distribution with non-integrable kernel $\propto \frac{1}{|x|}$ by substituting the derivative of a distribution with kernel $\propto \log(|y|)$, which is integrable.

In the above calculations, we introduced a mass/energy-scale $M$. It is now an important question to ask how the distribution changes under transformations of this scale, i.e. renormalization group (RG) transformations generated by $M\frac{\partial}{\partial M}$, so called RG flows. We will now show that it is only the part $const \cdot \delta(x)$, i.e. the part that is fixed by arbitrary renormalization constants that will change.

First of all, using $\partial_x \mathrm{sign}(x) = 2\delta(x)$, we see:

$$f'(x) = \frac{-1}{x} \theta\left(1 - M|x|\right) \mathrm{sign}(x) - \log(M|x|) \theta\left(1 - M|x|\right) 2\delta(x)$$

Then, we start with:

$$M\frac{\partial}{\partial M} T_{\frac{1}{|x|}}[\phi] = M\frac{\partial}{\partial M} \left( \int dx \frac{(1 - w(x))\phi(x)}{|x|} + T_{-\partial f}[\phi] \right)$$

Because of $\frac{1-w(x)}{|x|} = \frac{\theta(M|x|-1)}{|x|}$ in our example, the first term becomes a distribution with kernel

$$M\frac{\partial}{\partial M}\frac{\theta\left(M|x|-1\right)}{|x|} = M\delta(M|x|-1). \tag{56}$$

The second term in contrast becomes a distribution with kernel

$$M\frac{\partial}{\partial M}(-f'(x)) = -M\delta(M|x|-1) + M\frac{\partial}{\partial M}\left[2\log(M|x|)\theta\left(1-M|x|\right)\delta(x)\right].$$

The first term of this expression obviously cancels with the contribution from (56), so $M\frac{\partial}{\partial M}T_{\frac{1}{|x|}}$ turns out to be a distribution with kernel:

$$\begin{aligned}
&M\frac{\partial}{\partial M}\left[2\log(M|x|)\theta\left(1-M|x|\right)\delta(x)\right] \\
&= 2\delta(x)\left[\theta\left(1-M|x|\right) - \log(M|x|)M\delta\left(1-M|x|\right)|x|\right] \\
&= 2\delta(x)
\end{aligned}$$

In the last step, we used the presence of the factor $\delta(x)$ (under an integral!) to set $\log(M|x|)|x| = 0$ and $\theta\left(1-M|x|\right) = 1$. So, under a renormalization group transformation, the distribution changes by $\delta T \propto const \cdot \delta(x)$, that means that a change of energy-scale corresponds to a change of the (at the beginning) arbitrarily selected renormalization coefficients.

### 3.6 What we have achieved in this section

We have seen a way to recast what looks like divergent Feynman diagrams as to what looks like distributions for non-integrable functions. We could then turn these into proper distributions by first restricting the space of test-functions and then extend them to a full distribution, possibly at the price of a finite number of undetermined numerical constants. Those have to be determined by a finite number of measurements.

In order for the number of introduced parameters for all Feynman diagrams of the theory to be finite, the scaling degrees of all appearing distributions in all diagramms have to be below some maximum, otherwise the theory is not renormalizable.

## 4 Summary

The material in these notes will not be useful for any concrete calculation in quantum field theory that a physicist might be interested in. But they might give him or her some confidence that the calculation envisaged has a chance to be meaningful.

We tried to present material that is in no sense original but still is probably not covered in most introductions to quantum field theory. Hopefully, it helps to refute some of the prejudices against (perturbative) quantum field theory that mathematically minded people may have and helps others to better understand how far the hand waving arguments that we use in our daily work can carry.

In particular, we put our emphasis on two points: Even if the perturbative expansion is divergent as a power series it can serve two purposes: The first terms do

provide a numerically good approximation to the true, non-perturbative result and all terms taken together can indeed recover the full result but only in terms of Borel resummation rather than as a power series. Second, unphysical infinite momentum integrals in the computation of Feynman diagrams can be avoided when properly expressed in terms of distributions. The renormalization of coupling constants is then expressed as the problem to extend a distribution from a subspace to all test functions. The language of distribution theory allows one to avoid mathematically ill-defined divergent expressions altogether.

## Acknowledgements

A lot of the material presented we learned from Klaus Fredenhagen, Dirk Prange and Marcel Vonk. We would like to thank the Elitemasterprogramme "Theoretical and Mathematical Physics" and Elitenetzwerk Bayern.

## Bibliography

[1] H. Osborn, "Advanced quantum field theory lecture notes." available at http://www.damtp.cam.ac.uk/user/ho/Notes.pdf, April, 2007.

[2] F. J. Dyson, *Divergence of perturbation theory in quantum electodynamics*, *Phys.Rev.* **85** (1952) 631–632.

[3] M. Nakahara, *Geometry, Topology and Physics*. Taylor & Francis, 2003.

[4] M. Reed and B. Simon, *Analysis of operators*, vol. 4 of *Methods of modern mathematical physics*. Academic Press, 1978.

[5] G. Scharf, *Finite Quantum Electrodynamics: The Causal Approach*. Springer, New York, 1995.

[6] R. Brunetti and K. Fredenhagen, *Quantum field theory on curved backgrounds*, 0901.2063.

[7] D. Prange, *Epstein-glaser renormalization and differential renormalization*, *J.Phys.A* **A32** (1999) 2225–2238, [hep-th/9710225].